Overview of 18 comparisons in the MetaBLIND study for analysis (Ia) Blinding of patients in trials with patient reported outcomes.							
Review	**CD number**	**Analysis and outcome measure**	**Blinded studies**	**Unblinded and unclear studies**	**Weight**	**Ratio of odds ratios**	**Notes**
Anticonvulsants for alcohol dependence	CD008544	Analysis 6.2. Comparison 6 Anticonvulsants versus other medications (naltrexone), Outcome 2 Severe relapse, dichotomous outcome.	3	1	5.37	0.28	The four studies used different anticonvulsants in the intervention arm. Two trials used topiramate, one used pregabalin, and one used oxcarbazepine.
Clonidine premedication for postoperative analgesia in children	CD009633	Analysis 2.1. Comparison 2 Clonidine versus midazolam, Outcome 1 Number pain-free in PACU.	1	1	3.32	0.31	The unblinded study by Schmidt and colleagues had a third treatment arm where participants received dexmedetomidine. This might have reduced expectancy effects for clonidine.
Valproate (valproic acid or sodium valproate or a combination of the two) for the prophylaxis of episodic migraine in adults	CD010611	Analysis 1.1. Comparison 1 Divalproex sodium versus placebo, Outcome 1 ORs for responders (patients with ≥ 50% reduction in headache frequency).	1	3	4.62	0.31	Three studies were rated as probably blinded while the trial by Kaniecki and colleagues was rated as 'unclear'. Because trials with unclear blinding were grouped with unblinded trials, this comparison depends on whether participants in the trial by Kaniecki and colleagues were not blinded. The trial reports, however, that patients were blinded to the use of placebo. The Cochrane reviewers highlight that blinding could have been broken by the knowledge of therapists and differences in the appearance or taste of the intervention versus the placebo but this remains unclear. The fact that all four studies in this comparison report that patients were blinded, makes them a problematic choice to measure the effect of blinding trial participants. In contrast to the three other trials, the study by Kaniecki et al. also tested the use of propranolol hydrochloride in a 5-phases crossover design. This could have reduced expectancy effects of divalproex sodium.
Antibiotics for preventing complications in children with measles	CD001477	Analysis 1.2. Comparison 1 Antibiotic versus placebo or no antibiotic (excluding children with pneumonia or sepsis on admission), Outcome 2 Development of diarrhoea.	1	1	1.06	0.50	The unblinded trial by Karelitz and colleagues was conducted on children in New York in the 1950s and compared benzethacil or aqueous procaine penicillin with no treatment. The blinded trial by Garly and colleagues included patients of a measle outbreak in Guinea-Bissau in 1998 and compared sulfamethoxazole-trimethoprim with a placebo. These differences complicate a direct comparison.

Short-term treatment with proton pump inhibitors, H2-receptor antagonists and prokinetics for gastro-oesophageal reflux disease-like symptoms and endoscopy negative reflux disease	CD002095	Analysis 4.1. Comparison 4 PPI versus H2RA, Outcome 1 Heartburn remission 4.1.2 Endoscopy negative reflux disease	2	2	6.73	0.56	The four trials used different drugs in the control group: nizatidine, ranitidine, famotidine, or cimetidine. While three trials used omeprazole as the intervention, one study (Armstrong et al. 2001) used pantoprazole. Despite a description of being “double-blind”, blinding of participants of the trial by Bate and colleagues was rated as ‘unclear’. Therefore it was classified in the group of unblinded trials. The trial by Fujiwara and colleagues was rated as probably not blinded. The study did not provide information on blinding of participants.
Megestrol acetate for treatment of anorexia-cachexia syndrome	CD004310	Analysis 1.1. Comparison 1 Megestrol acetate versus placebo (ITT), Outcome 1 Appetite improvement	2	3	5.66	0.58	Four trials included patients with cancer while the trial by Von Roenn and colleagues recruited patients with AIDS. Three trials were rated as probably blinded while the blinding status of the two studies by Schmoll and colleagues were rated as ‘unclear’. Although blinding is not explicitly described in these studies, the fact that they used a placebo-control design suggests that participants may have been blinded to treatment allocation. This comparison, therefore, lacks a trial where participants were clearly not blinded.
Antidepressants for smoking cessation	CD000031	Analysis 1.1. Comparison 1 Bupropion. Abstinence at 6m or greater follow-up, Outcome 1 Bupropion versus placebo/control. Subgroups by length of follow-up.	1	1	6.98	0.64	The blinded trial by Planer and colleagues compared bupropion against placebo. In the unblinded trial by Wittchen and colleagues, there were four treatment arms: minimal intervention, cognitive behavior therapy (CBT), CBT + bupropion, and CBT + nicotine replacement therapy. This study design might have weakened perceptions of CBT as the control group and CBT + bupropion as the intervention group. While the study by Wittchen et al. tested the use of bupropion in regular smokers in primary care, the study by Planer et al. tested bupropion in smokers hospitalized with acute coronary syndrome.
Interventions to promote informed consent for patients undergoing surgical	CD009445	Analysis 1.5. Comparison 1 All studies: Interventions that promote informed consent, Outcome 5	25	1	4.92	0.68	This comparison depends on the trial by Garden and colleagues which was the only one rated as blinded. This trial is problematic because it did not test the effectiveness of an intervention to treat a medical condition. Instead, it tested the knowledge recall of

and other invasive healthcare procedures		Knowledge/retention/recall - immediate: continuous data.					participants who were given different types of information sheets. As the authors indicate the information provided in the full information sheet (the intervention) was more relevant to the questionnaire used as the outcome measure than the standard leaflet (the control). They write: "The 'standard' leaflet had been constructed without reference to this study, by consensus between a large number of specialist anaesthetists. In contrast, the knowledge questionnaires and the 'full' information sheet were designed by the investigators to address information thought important. It turned out that the information in the 'standard' leaflet (unlike the 'full' leaflet) did not actually cover all the issues addressed by the knowledge questionnaire." In addition, the trials in this comparison used different Interventions to promote informed consent (information leaflets, a computer-based multimedia presentation, repeat-backs, teach-the-teacher method, etc.) and different control groups, which complicate a direct comparison.
Strategies for partner notification for sexually transmitted infections, including HIV	CD002843	Analysis 1.3. Comparison 1 Enhanced patient referral versus simple patient referral, Outcome 3 Number of partners notified.	3	1	9.26	0.83	This comparison depends on the trial by Østergaard and colleagues which was the only one that was rated as blinded. In this trial, the number of partners notified was not an outcome measure. The trial investigated two strategies – home sampling versus office sampling – to encourage previous sex partners of index patients to get tested for Chlamydia trachomatis. Given that index patients were blinded to the sampling method and that they only had to pass on the test kit to their previous sex partners, the number of partners they notified is not dependent on the intervention.
Surgical management of pelvic organ prolapse in women	CD004014	Analysis 2.1. Comparison 2 One method of anterior prolapse repair versus another surgical method, Outcome 1 Number of women with prolapse symptoms (subjective failure).	3	1	7.9	0.84	The trials compared two surgical interventions. Expectancy effects might have been reduced as it might not have been clear to participants which was the intervention group and which the control group.

		9 anterior colporrhaphy versus any transvaginal polypropylene mesh					
Antibiotics for sore throat	CD000023	Analysis 1.1. Comparison 1 Antibiotics versus placebo for the treatment of sore throats: symptom of sore throat, Outcome 1Symptom of sore throat on day 3.	6	9	7.87	1.16	5 of the 6 studies in the unblinded/unclear group were conducted in the 1950s. At the time, explicit descriptions of the blinding process might have been less common. Three of these trials (Denny 1953, Chapple 1956, and Landsman 1951) do not mention blinding but used a placebo control group, suggesting that participants might have been blinded. Inclusion criteria for this meta-analysis were broad, namely "patients presenting to primary care facilities with symptoms of sore throat." This might have resulted in significant differences in patient samples. 3 unblinded/unclear studies (Brink 1951, Denny 1953, and Brumfit 1957) and 1 blinded study (McDonald 1951) recruited soldiers in the army. 3 unblinded/unclear studies (Little 1997, Chapple 1956, and Landsman 1951) and 5 blinded studies (De Meyere 1992, Dagnelie 1996, Peterson 1997, Whitfield 1981, and Zwart 2000) were conducted in general practice. Three blinded studies included children or young adults (Middleton 1988, El-Daher 1991, and Zwart 2003). The 3 studies that reported the largest effects (Middleton 1988, De Meyere 1992, and El-Daher 1991 were blinded but included only patients who were positive for Group A streptococci. The Cochrane review found evidence that the effectiveness of antibiotics for sore throat is larger in this subgroup than in those negative for Group A streptococci. In 3 unblinded studies (Brink 1951, Denny 1953, and Brumfit 1957) the majority of the extracted data came from participants who were positive for Group A streptococci. One blinded study (Peterson 1997) only included patients who were negative for Group A streptococci. The blinded trial by Middleton and colleagues found the largest effect but it reported on the number of patients whose symptoms improved after 48 hours, rather than

							the number of patients that no longer had symptoms after 3 days.
Music for stress and anxiety reduction in coronary heart disease patients	CD006577	Analysis 1.1. Comparison 1 Music versus standard care, Outcome 1 Psychological distress.	4	1	3.25	1.37	The comparison depends on the trial by Karin Schou which was the only one rated as blinded. This study was an unpublished Ph.D. thesis with data from only 17 participants. Given the difference in contact time with therapists, it is highly unlikely that this trial was blinded. While the control group had scheduled rest on their own, the intervention consisted of: "a music therapy treatment, consisting of a receptive music therapy method Guided Relaxation with Music, (Music therapy and medicine) in which the participants received personalised individual sessions of guided relaxation with music with a trained music therapy research team member (RTM) in the role of guiding a body relaxation based on the patient's preferred style of music for relaxation." The Cochrane review rated the trial by Schou at low risk of performance bias but it adds that "music therapist and participants could not be blinded given the interactive nature of the music therapy session." Elsewhere the review notes that "since participants cannot be blinded in a music intervention trial, we did not downgrade studies for not blinding the participants."
Hormone therapy for sexual function in perimenopausal and postmenopausal women	CD009672	Analysis 2.1. Comparison 2 Estrogens + progestogens, Outcome 1 Composite score.	2	1	8.97	1.42	This comparison depends on the WISDOM study which was the only one rated as blinded. This trial found a minor effect in favor of the intervention (standardized mean difference of 0.11) but it also suffered from incomplete outcome data. At one year follow-up 1043 participants in the intervention and 1087 in the control group completed the items of the women's health questionnaire. The only exception is the sexual functioning domain from which data was extracted. Here, only 588 in the intervention and 569 in the control group filled out the questionnaire. This difference was not explained. Therefore, the Cochrane review rated the trial at high risk of attrition bias. The intervention differs in the three trials: estradiol and dienogest, oestrogen and medroxyprogesterone, or 17-β-estradiol and norethisterone acetate. The trials also

							used different outcome measures: Rosen's female sexual function index, Subjective Symptoms Assessment Profile, and the sexual functioning subscale of the Women's Health Questionnaire. The study by Osmanagaoglu and colleagues included a third arm where patients received tibolone. This might have reduced expectancy effects for estradiol + dienogest. The study by Czarnecka is only available in Polish and was rated at high risk of bias due to baseline differences for several outcomes.
Neuropsychological rehabilitation for multiple sclerosis	CD009131	Analysis 1.11. Comparison 1 Cognitive training versus any control, Outcome 11 Quality of life	2	1	5.02	1.58	There were minor differences in the intervention and control arm, length of treatment, inclusion criteria, and the outcome measure used across the three trials. The blinded trial by Solari and colleagues found an effect in favor of the intervention at the 8-week assessment (post-treatment) on the mental health composite of the MS Quality of Life Questionnaire (MSQOL). At the 16-week assessment, however, the control group did better than the intervention group on this outcome. Similarly, on the cognitive function subscale of MSQOL, the control group outperformed the intervention group at both the 8- and 16-week assessments. Data on the physical health composite of MSQOL was not reported. The authors of the study concluded: "This trial does not support the efficacy of specific memory and attention retraining in MS." The blinded trial by Solari and colleagues was rated as high risk of bias because of a "significant difference between intervention and control groups in age." The unblinded trial by Vogt and colleagues reported better quality of life in the control group (122.93±32.27) than the intervention group (118.60±34.08) post-treatment. This difference, however, was smaller than the difference at baseline between the control group (125.99±32.56) and the intervention group (117.23±29.41).

Methotrexate for ankylosing spondylitis	CD004524	Analysis 1.12. Comparison 1 MTX versus No MTX, Outcome 12 Patient global assessment (different scales, the more severe the disease, the higher the score) (end point value).	1	1	3.99	1.64	The blinded trial by Gonzalez-Lopez and colleagues compared Methotrexate (MTX) with a placebo while the open-label study by Altan and colleagues compared MTX + naproxen versus naproxen alone.
Decision aids for people facing health treatment or screening decisions	CD001431	Analysis 1.1. Comparison 1 Knowledge, Outcome 1 Knowledge: DA vs usual care - all studies	40	2	10.86	1.93	This comparison does not test the effectiveness of an intervention to treat a medical condition but the knowledge recall of participants who were given different types of decision aids. Only two trials were rated as blinded: Shorten 2005 and Steckelberg 2011. In both trials, the decision aid used as the intervention had more relevant information to the knowledge questionnaire than the information provided to the control group. In the trial by Steckelberg for example the control group received a standard information leaflet where “no quantitative information on individual risk or benefit is included, and harm is incompletely communicated.” In the trial by Shorten and colleagues participants in the control group received no information leaflet but routine pregnancy care. Knowledge was assessed with a questionnaire that was developed for the study based on key risk and benefit information contained in the decision-aid. While participants in the intervention group of Shorten 2005 and Steckelberg 2011 scored better than the control group on the knowledge questionnaire, this did not result in a change in the uptake of cesarean section and colorectal cancer screening respectively. Both trials, therefore, concluded that the intervention was ineffective. The decision aids compared in this analysis differed significantly: some trials used an information leaflet, others an audiotape, video website, or an interactive computer program. The target population and medical decisions also differed considerably, from people considering a screening test for Down syndrome, revascularization surgery, feeding tube placement, cesarean section, prostate cancer screening, etc. Therefore these trials may not be directly comparable.

Pharmacological interventions for pruritus in adult palliative care patients	CD008320	Comparison 5. Rifampin or rifampicin versus placebo or standard medication Pruritus score	1	2	1.03	3.3	In contrast to the two blinded trials in this analysis, the unblinded trial by Bachs and colleagues did not compare rifampicin to a placebo in a cross-over design but to another drug called phenobarbitone in a parallel design. This might have reduced the efficacy of rifampicin. For a relevant comparison, the unblinded trial by Bachs and colleagues should have compared to a placebo as well. The trials by Ghent et al. and Podesta et al. were rated as probably blinded even though there was a high risk that blinding was broken due to a darkening of the urine by rifampicin. As noted by Podesta et al.: "Most patients taking rifampin develop a red-orange coloration of the urine, thus allowing them to identify the experimental period, and eliminating the double-blind nature of this study." Ghent and colleagues noted: "As rifampin discolors urine, there was a potential for determining the treatment being given, and eliminating the double blind nature of this study."
Paracervical local anaesthesia for cervical dilatation and uterine intervention	CD005056	Analysis 3.1. Comparison 3 Paracervical block versus other regional anaesthesia, Outcome 1 Pain during cervical dilatation.	1	1	3.54	36.89	The blinded study by Mankowski and colleagues was rated as high quality. It found no difference between paracervical block and intracervical block. In contrast, the smaller and low-quality trial by Yacizi found a large difference (a standardized mean difference of 2.08) in favor of paracervical block. Moustgaard and colleagues viewed intracervical block as the intervention and paracervical block as the control condition. Therefore this comparison was given a ratio of odds ratios that speaks strongly against an effect of blinding. This is questionable because the trial by Yacizi also included a placebo group. The study compared two forms of anesthesia with a placebo and a fourth group where both interventions were combined. To participants, it might not have been clear that intracervical block was the intervention and paracervical block the control condition. The fact that the unblinded trial found a much stronger effect than the blinded trial, could support rather than contradict the importance of blinding.